A NUMERICAL TAXONOMIC ANALYSIS OF INTERSPECIFIC MORPHOLOGICAL DIFFERENCES IN TWO CLOSELY RELATED SPECIES OF *CICADA* (HOMOPTERA, CICADIDAE) IN PORTUGAL

J. A. Quartau[1]

ABSTRACT.—*Cicada orni* Linnaeus is among the most common and widespread cicadas in Portugal, and, unless a critical study of the male genitalia is made, it is easily confused with the much less widely distributed *C. barbara lusitanica* Boulard. These species are morphologically very similar and sometimes difficult to separate using existing keys. This study attempts to test the discriminating capabilities of numerical techniques commonly used for classificatory purposes, as well as to discover the most effective characters to distinguish between the two species. For these purposes, cluster analysis and principal component analysis were applied to a sample of 64 male specimens characterized by 40 characters (33 derived from the external morphology and 7 from genitalia). In WPGMA cluster analysis, product-moment correlations gave a better separation between these species than did taxonomic distance coefficients; moreover, the analysis derived from the genital characters alone gave better separation than the analyses based on the 33 external characters. Principal component analysis yielded a clear, interspecific separation along the first axis. The best characters to discriminate between males of the two species were the lengths of the pygofer (and its dorsal spine), the tenth abdominal segment, and the appendages of the latter (which are smaller in *barbara lusitanica*), as well as the width of the shaft of the aedeagus (thinner in *orni*). Finally, the uniformity of the general clustering pattern resulting from the two multivariate techniques suggests the presence of two distinct species, as also clearly indicated by behavioral data.

Cicada orni Linnaeus is among the most common and widespread cicadas in Portugal, and, unless a critical study of the male genitalia is made, it is easily confused with the much less widely distributed *C. barbara lusitanica* Boulard (Quartau and Fonseca 1988). As live specimens, however, they are easily distinguished by the male calling songs, which are quite distinct. Oscillograms are found in Claridge et al. (1979) and Boulard (1982), respectively, for *C. orni* and *C. barbara lusitanica*.

The two species are externally very similar and sometimes even difficult to separate by existing keys (e.g., Gómez-Menor 1957). In fact, the main distinguishing character used for their separation has been the presence in *barbara* of only two spots on the cross-veins of the forewings instead of four; however, some specimens of *barbara lusitanica* have the full four spots as they occur in *orni* (Fig. 4).

Boulard (1982), when describing the Portuguese form of *C. barbara*, which he originally named *lusitanica*, provided a good diagnosis of the genital characters of this species. However, no detailed comparison of the two species has been made, nor has any type of multiple-character analysis involving the simultaneous use of several measurements or counts been attempted. It was felt of interest, therefore, to see how far some common techniques of numerical taxonomy would discriminate between this pair of closely related species.

This study was undertaken with two main objectives in mind. The first was to apply current techniques of numerical taxonomy commonly used for classificatory purposes with the aim of testing their general discriminating power with respect to these two species. The techniques chosen were a form of hierarchical cluster analysis and principal component analysis. It is known that apart from the explicit use of the former, principal component analysis can also serve as a cluster technique of great generality and can be used to distinguish pairs of putative morphs as in the classical study of Temple (1968). The second objective was to discover new characters that might help to separate *C. orni* from *C. barbara*.

[1]Depto. de Zoologia e Antropologia, Faculdade de Ciências de Lisboa, 1200 Lisboa, Portugal.

MATERIAL

The data on which this study is based were taken from dried male specimens (OTUs) of 32 *Cicada orni* and 32 *C. barbara lusitanica* (Table 1). These samples were mostly taken by the author in Portugal: all 32 males of *orni* were collected in central Portugal; an equal number of males of *barbara lusitanica* were taken in several areas of Algarve (the southern province of Portugal), where the species appears to be particularly common, with the exception of two specimens only that were collected in Sesimbra (south of Lisbon). The localities and sample sizes are *C. orni:* Alburitel, Vila Nova de Ourém (n = 32); *C. barbara lusitanica:* Carvoeiro (n = 25), Praia da Rocha (n = 4), Serra de Monchique (n = 1), and Sesimbra (n = 2).

TABLE 1. List of specimens (males) of *Cicada orni* Linnaeus and of *C. barbara lusitanica* Boulard investigated (OTUs).

OTUs	Locality, date of capture, and collector
Cicada orni	
1–15	Alburitel, 10.viii.1979, J. A. Quartau
16–32	Alburitel, vii.1971, J. A. Quartau
C. barbara lusitanica	
33–34	Carvoeiro, 14.viii.1966, P. D. Rodrigues
35–38	Praia da Rocha, 17.viii.1973, J. A. Quartau
39–40	Carvoeiro, 14.vii.1978, J. A. Quartau
41–47	Carvoeiro, 30.vii.1978, J. A. Quartau
48–55	Carvoeiro, 28.vii.1978, J. A. Quartau
56–57	Carvoeiro, 9.viii-10.ix.1980, L. Mendes
58–59	Sesimbra, 2.viii.1980, J. A. Quartau
60	Carvoeiro, 31.vii.1978, J. A. Quartau
61–62	Carvoeiro, 18.vii.1978, J. A. Quartau
63	Monchique, 2.ix.1971, F. Carvalho
64	Carvoeiro, 24.viii.1981, J. A. Quartau

METHODS

Measurements and Counts

Thirty-seven of the 40 characters were measurements; the remaining 3 were counts. Measurements were made using a Wild M3 microscope with a graduated eyepiece and were taken as described in Table 2 or as illustrated in Figures 1–9. Of these 40 characters, 33 refer to external morphology and the remaining 7 to male genitalia.

Data Analysis

Data processing was carried out on the CDC 6500 computer at the Imperial College Computer Center (University of London) using two multivariate statistical programs developed by Prof. R. G. Davies (Department of Pure and Applied Biology, Imperial College) for cluster analysis and ordination (Quartau and Davies 1983, 1985).

In most analyses, characters were standardized by expressing each state as a deviation from the mean in standard deviation units.

For Q-mode analysis, taxonomic distances as well as product-moment correlations were found and structured by the WPGMA method of cluster analysis (Weighted Pair Group Method with Arithmetic Averaging). Phenograms, expressing the phenetic relationships among the OTUs in a hierarchy of increasingly larger clusters, were thus obtained (Figs. 10–16).

For R-mode analysis, character correlations based on data standardized by OTUs were subjected to principal component analysis (PCA). This ordination method transforms the original characters, generally continuous, correlated characters, into a suite of uncorrelated, composite variables—the principal components (principal axes). In addition to being mutually independent, these components account for maximum variance as follows: the variance along the first axis (i.e., the corresponding eigenvalue) is the maximum possible. The second axis describes the next largest variance orthogonal to (uncorrelated with) the first. The third axis follows similarly but is independent of both first two axes, and so on, for as many axes as one wishes to extract (e.g., Gibson et al. 1984). A transposed matrix of the character loadings was post-multiplied by the standardized data matrix to yield a matrix of OTU projections in the principal component space. Two-dimensional ordination diagrams of the representations of the two species, together with the character loadings (scaled eigenvectors), were thus obtained (Figs. 17–18, Table 3).

RESULTS

Phenograms

The seven phenograms resulting from various analyses based on all characters, on the genitalia only, or on the external characters alone are shown in Figures 10–16. The WPGMA clustering technique was followed

TABLE 2. Description of characters: measurements and counts (terminology mostly follows Myers [1928]).

Character No.	Description
1.	Overall length measured from tip of crown to apical margin of the right forewing with the latter in position of rest alongside the body (Fig. 1).
2.	Length of crown measured along a medial line passing through the median ocellus (Fig. 1).
3.	Minimum distance between the ocular sutures measured along the paired ocelli (Fig. 2).
4.	Medial length of frons measured dorsally as indicated (Fig. 2).
5.	Medial length of pronotum measured dorsally as indicated (Fig. 1).
6.	Medial length of mesonotum measured dorsally from anterior margin to posterior margin of cruciform elevation or scutellum (Fig. 1).
7.	Width of pronotum measured at the level of anterior lateral margins (Fig. 1).
8.	Width of pronotum measured at the level of posterolateral margins (Fig. 1).
9.	Width of crown measured at the level of median ocellus and as indicated (Fig. 2).
10.	Inner distance between the paired ocelli (Fig. 2).
11.	Distance between the right paired ocelli and the right ocular suture as indicated (Fig. 2).
12.	Distance between the base of the left antenna and the left ocular suture as indicated (Fig. 3).
13.	Inner distance between the base of antennae (Fig. 3).
14.	Length of frons as illustrated (Fig. 3).
15.	Length of clypeus as illustrated (Fig. 3).
16.	Length of the exposed part of beak.
17.	Length of dorsal margin of the left fore femur as illustrated (Fig. 5).
18.	Length of ventral margin of the left fore femur as illustrated (Fig. 5).
19.	Length of basal spine in ventral margin of left fore femur as illustrated (Fig. 5).
20.	Length of apical spine in ventral margin of left fore femur as illustrated (Fig. 5).
21.	Distance between tips of the apical and basal spines in the ventral margin of the left fore femur as illustrated (Fig. 5).
22.	Distance from anterior right corner to posterior left corner of left operculum as illustrated (Fig. 6).
23.	Distance from anterior left corner to posterior right corner of left operculum as illustrated (Fig. 6).
24.	Length of right forewing as illustrated (Fig. 1).
25.	Greatest width of right forewing as illustrated (Fig. 4).
26.	Length of subcostal cell ("gancho" cell of Gómez-Menor 1957) in right forewing (Fig. 4).
27.	Length of anterior margin of basal cell in right forewing (Fig. 4).
28.	Length of posterior margin of basal cell in right forewing (Fig. 4).
29.	Maximum width of basal cell in right forewing.
30.	Minimum width of basal cell in right forewing.
31.	Number of apical cells in right forewing.
32.	Number of cells other than apicals of right forewing.
33.	Number of spots in cross-veins of right forewing.
34.	Length of pygofer in lateral view as indicated (Fig. 7).
35.	Overall length of tenth abdominal segment as indicated (Fig. 7).
36.	Overall length of appendages of tenth abdominal segment as indicated (Fig. 7).
37.	Distance in basal curvature of shaft of aedeagus as indicated (Fig. 9).
38.	Width of shaft of aedeagus as indicated (Fig. 9).
39.	Width of shaft of aedeagus in area of curvature as illustrated (Fig. 9).
40.	Medial length of eighth sternite or hypandrium (Fig. 8).

in all; and Pearson's product-moment coefficient and the taxonomic distance coefficient were used as measures of taxonomic proximity.

(a) Genital analyses

Figures 10 and 11 illustrate a correlation and a distance phenogram, respectively, both based on the seven standardized variables of the male genitalia. Both analyses, notwithstanding their being based on a small number of variables, resulted in two main clusters, one with *C. orni* and the other with *C. barbara lusitanica*. However, in the former phenogram, the cluster of *barbara lusitanica* includes one specimen of *orni* (No. 13).

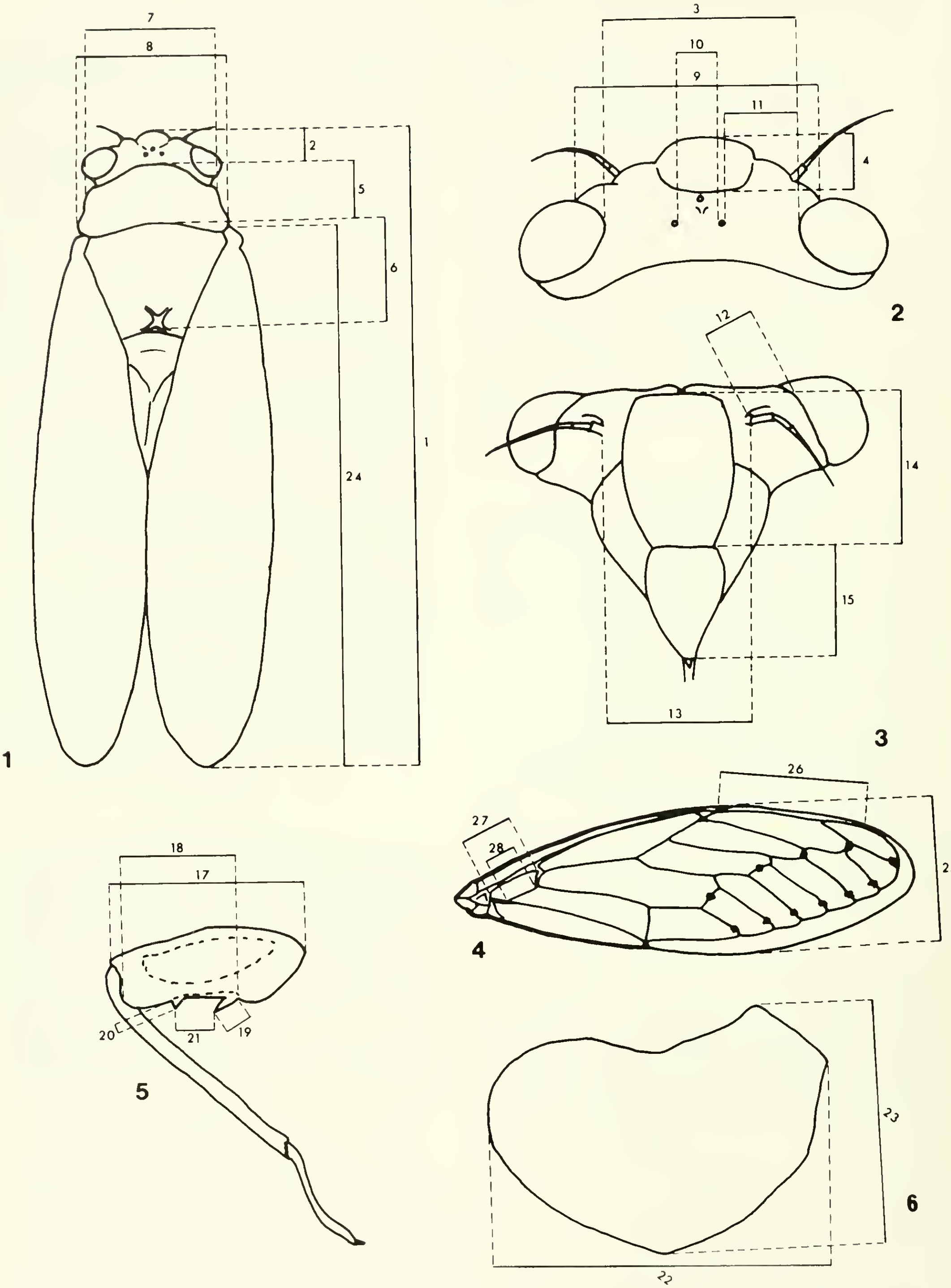
(b) External characters

These analyses resulted in the production of the phenograms depicted in Figures 12 and 13. Both were based on standardized data, but only the correlation coefficient succeeded in giving an almost complete separation of the two species of cicadas. In fact, OTUs were grouped into two main clusters as in the genital analyses, but specimen No. 19 belonging to *C. orni* appeared misplaced within *C. barbara lusitanica* (Fig. 12). On the contrary, the distance phenogram provided much less satisfactory results than the previous analysis, since each of the two major clusters incorporates elements of both species of cicadas (Fig. 13).

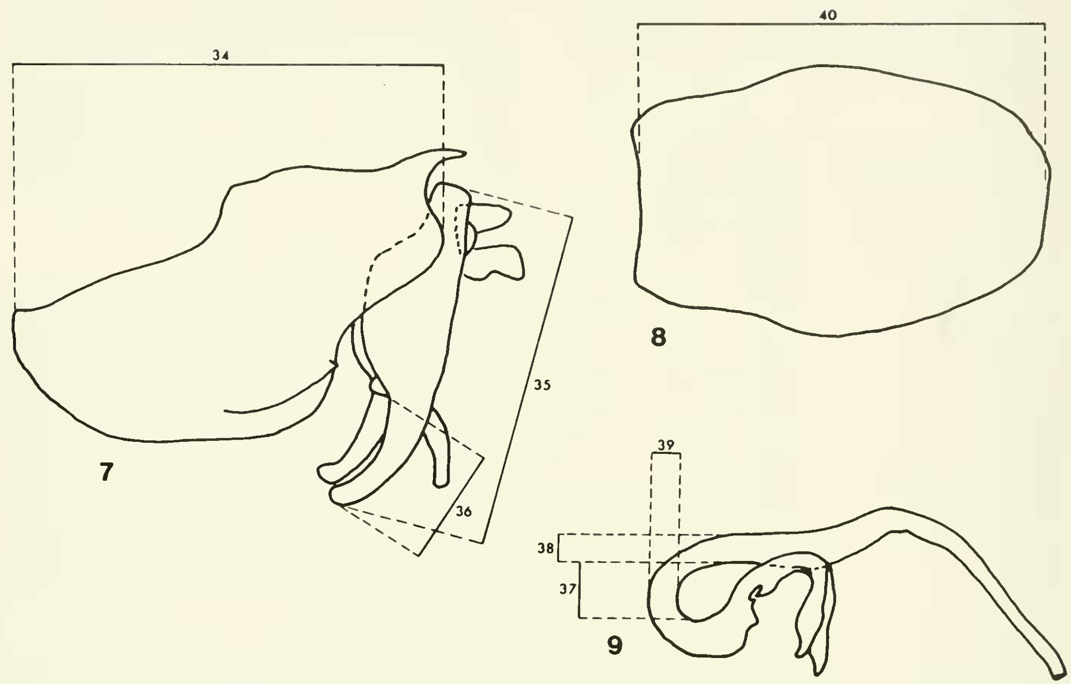
(c) Combined characters

The phenograms of this group of analyses, involving all 40 characters combined, are illustrated in Figures 14–16.

Considering the phenograms based on standardized data (Figs. 14, 16), it is clear that the correlation phenogram gave a much better distinction between the two species than the distance analysis. In fact, the latter (Fig. 16) clustered six specimens of *C. orni* with *C.*



Figs. 1–6. Diagrams of a male *Cicada orni* Linnaeus illustrating most of the measurements taken: 1, body, dorsal view; 2, head, dorsal view; 3, head, anterior view; 4, right forewing, dorsal view; 5, left foreleg, inner view; 6, left operculum, ventral view.



Figs. 7–9. As in Figures 1–6: 7, pygofer and tenth abdominal segment, ventral view; 8, eighth sternite or hypandrium, ventral view; 9, aedeagus, lateral view.

barbara lusitanica. Even when the data were unstandardized, correlations gave a good picture of the relationships between these two species (Fig. 15).

Principal Component Analysis

This analysis involved all 40 characters and was computed from a between-character correlation matrix based on data standardized by OTUs.

As in similar analyses carried out with leafhoppers of the genus *Batracomorphus* (Quartau 1983), slightly more than half (54%) of the total variation in the study was explained by the first three axes.

The first component accounted for 38.90% of the variation in the data and is interpreted as a contrast between the lengths of the pygofer, tenth abdominal segment, or appendages of the latter and the width of the shaft of the aedeagus. It does not represent overall size as commonly is the case, since many of the characters (Table 3) are not positively correlated with it (e.g., Jolicoeur and Mosimann 1960, Blackith and Reyment 1971, Baker 1980, Gibson et al. 1984, Shea 1985). In fact, it must represent both size and shape as has been pointed out by several authors (Mosimann 1970, Oxnard 1978, Humphries et al. 1981). A complete separation of *C. orni* and *C. barbara lusitanica* was given by the discrimination afforded by this axis, which is probably close to the orientation of the optimum discriminant function. The characters loading most heavily on this component (Table 3) are therefore of considerable taxonomic interest, since they are diagnostic for this pair of species. The highest negative scores, in decreasing order, were for characters numbered 35 (length of tenth abdominal segment), 34 (length of pygofer), and 36 (length of appendages of tenth abdominal segment). The highest positive score was for character numbered 38 (width of shaft of aedeagus).

The second principal component accounted for 8.51% of the total variation and was interpreted as a contrast between the number of spots in cross-veins of the wings and the width of the crown. It was most heavily loaded, negatively and positively, on characters numbered 33 and 9, respectively.

The third principal component accounted for 6.61% of the total variation and was interpreted as a factor resulting from the lengths of

TABLE 3. Eigenvector matrix (character loadings) in a principal component analysis of the matrix of correlations among the 40 variables (data standardized by OTUs.)

Variables		Scaled eigenvectors		
		I	II	III
1		0.296	−0.158	0.441
2		0.404	0.123	0.155
3		0.531	0.557	−0.174
4		0.843	−0.160	0.126
5		0.492	0.300	−0.119
6		0.037	−0.096	0.280
7		0.567	0.280	−0.069
8		0.314	0.365	−0.131
9		0.336	0.633	−0.133
10		0.738	−0.237	−0.000
11		0.687	0.155	−0.200
12		0.859	−0.127	−0.142
13		0.313	0.423	0.414
14		0.117	0.422	0.306
15		−0.604	−0.196	−0.120
16		−0.769	0.104	−0.192
17		−0.316	0.338	0.524
18		−0.667	0.174	0.375
19		0.565	0.252	−0.091
20		0.531	−0.066	0.111
21		−0.132	0.019	0.257
22		0.784	0.203	0.160
23		0.847	0.112	0.000
24		0.524	−0.340	0.060
25		0.683	−0.104	0.154
26		0.187	−0.207	−0.412
27		−0.334	0.126	−0.561
28		0.133	0.117	−0.549
29		−0.011	−0.207	0.471
30		0.129	−0.287	0.511
31		0.818	−0.490	−0.049
32		0.818	−0.479	−0.059
33		0.266	−0.745	−0.163
34		−0.952	−0.011	−0.040
35		−0.958	−0.153	−0.001
36		−0.944	−0.180	0.043
37		−0.869	−0.250	0.028
38		0.942	−0.065	−0.023
39		0.892	−0.249	0.024
40		−0.858	−0.125	−0.090
Latent roots		15.559	3.406	2.644
Percentage of variation	component	38.898	8.514	6.609
	cumulative	38.898	47.412	54.021

the anterior and posterior margins of the basal cell of the wings. It was most heavily loaded (negative scores) on characters numbered 27 and 28.

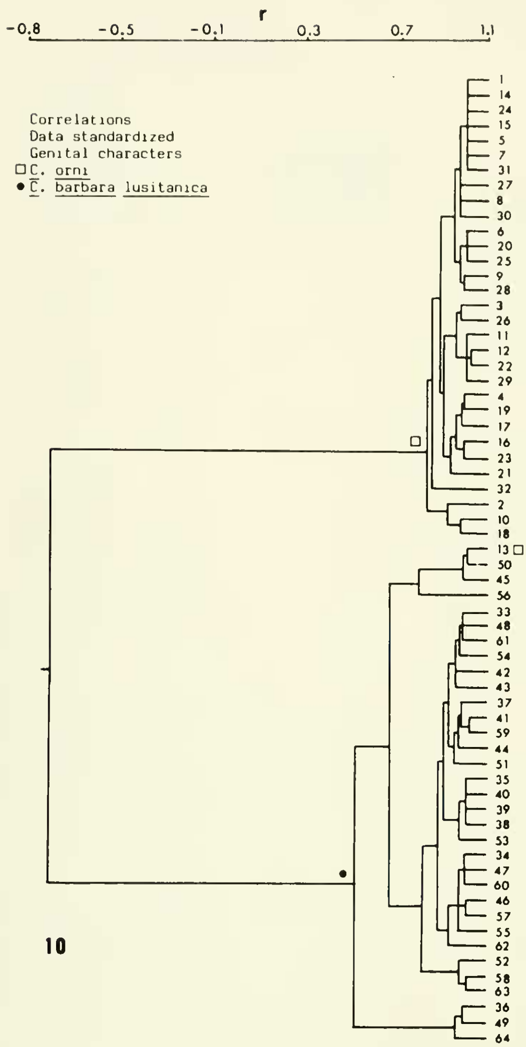
Neither the second nor the third axis leads to a separation of the two species of cicadas. Figures 17 and 18 are two-dimensional views of the relationships among specimens of both species in the space determined by component I combined with component II and by component I with component III, respectively. It is clear that these two plots gave a good separation between *C. orni* and *C. barbara lusitanica*. The plot combining axes II and III did not, however, succeed in giving a correct assignment of the two species and therefore was not illustrated here.

CONCLUSIONS

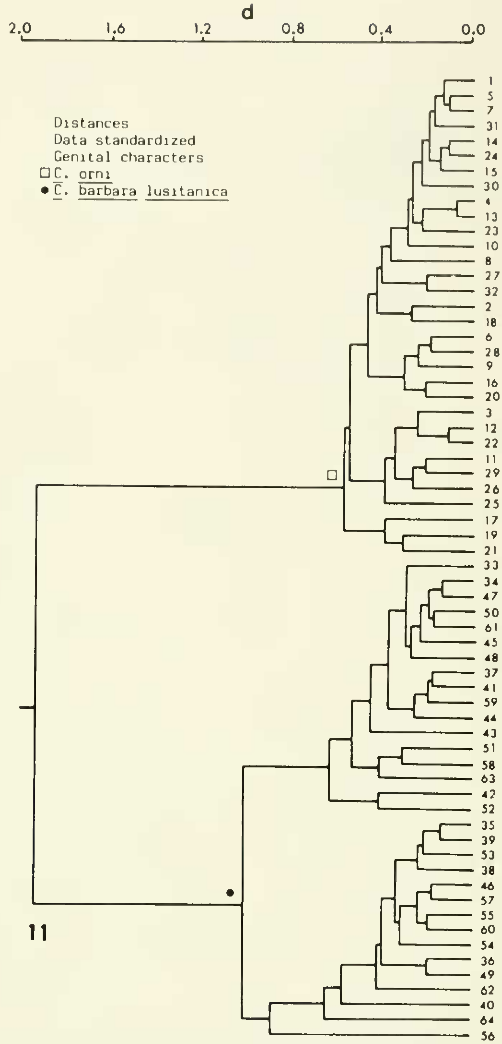
Cluster analysis and principal component analysis are two of the methods most commonly used for recognition of group structure in numerical taxonomy (e.g., Quartau 1987). This study suggests that the two general methods are also very useful for discriminating between pairs of closely related species. In fact, because of the uniformity of the general clustering pattern that resulted from the application of both methods, it is clear that two distinct species exist, a fact also indicated by behavioral data.

However, concerning cluster analysis, it is worthwhile noting that the hierarchical structure within each of the two major groups of OTUs differs a good deal from one particular method to another. Moreover, it appeared that correlations were more effective than taxonomic distances in describing relationships between the two cicadas, a result in keeping with Boyce (1964) or Cheetam (1968), for instance. It is interesting to note, in this regard, that such finding is in disagreement with a similar study carried out with leafhoppers (Quartau and Davies 1983) or with results based on other groups (e.g., Smith 1972). Finally, the failure of the cluster analysis using distances with nongenital characters shows that the use of such techniques needs to be undertaken with care and that it might be best to use a consensus of several clustering techniques when applying numerical methods to a novel taxonomic situation.

Principal component analysis succeeded in giving a good distinction of the two species along axis I, in spite of its accounting for only 38.90% of the total variation. This analysis also showed that the main distinguishing characters between *C. orni* and *C. barbara lusitanica* are the following: lengths of the pygofer, the tenth abdominal segment, and the appendages of the latter (which are smaller in *C. barbara lusitanica*, Figs. 19, 22), as well as the width of the shaft of the aedeagus (which is



10



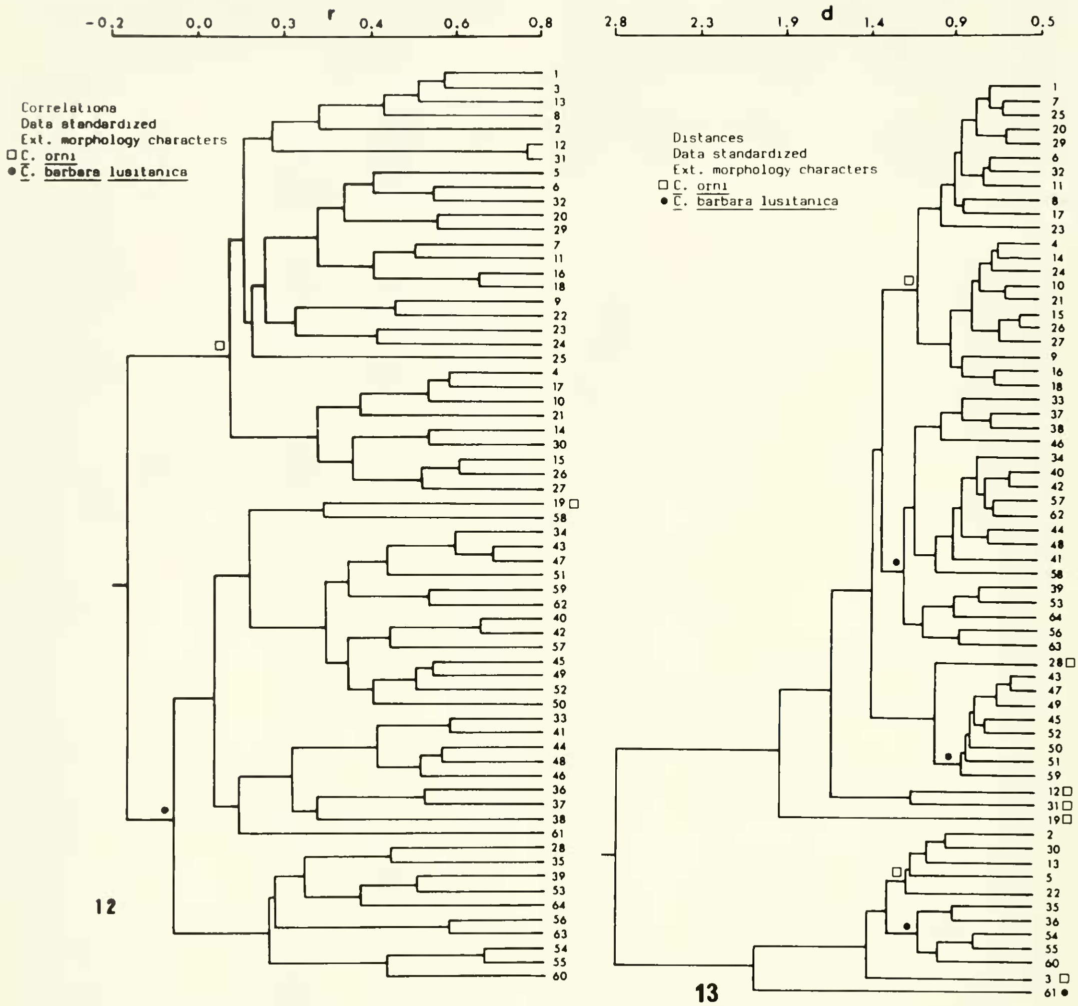
11

Figs. 10–11. 10, correlation phenogram based on the seven genital characters with standardized data; 11, distance phenogram based on the seven genital characters with standardized data.

thinner in *orni*, Figs. 20, 23). Moreover, detailed examination of the male genitalia also showed that the dorsal spine of the pygofer is smaller in *barbara lusitanica* than in *orni* (Figs. 19, 22).

Acknowledgments

I thank Prof. R. G. Davies, Department of Pure and Applied Biology, Imperial College, for having made his computer programs available for my use as well as for his critical comments on the manuscript. Prof. R. M. Anderson is also acknowledged for having granted working facilities at the Imperial College (University of London). I am also grateful to my colleague Dr. M. L. Mathias for her help with the preparation of most phenograms, and to Miss A. R. Cruz for help with other illustrations and the measurements. For enabling me to publish this study in honor of Prof. P. Oman (Oregon State University, USA), who has been a source of encouragement since the beginning of my taxonomic studies on leafhoppers, I also acknowledge Dr. M. W. Nielson (Brigham Young University, USA) and Dr. H. D. Blocker (Kansas State University, USA).



Figs. 12–13. 12, correlation phenogram based on the 33 external morphological characters with standardized data; 13, distance phenogram based on the 33 external morphological characters with standardized data.

This paper was completed at Imperial College (University of London) under an INIC (Ministério da Educação, Lisbon) sabbatical fellowship.

LITERATURE CITED

BAKER, A. J. 1980. Morphometric differentiation in New Zealand populations of the house sparrow (*Passer domesticus*). Evolution 34: 638–653.

BLACKITH, R. E., AND R. A. REYMENT. 1971. Multivariate morphometrics. Academic Press, London.

BOULARD, M. 1982. Les cigales du Portugal, contribution a leur étude (Homoptera, Cicadidae). Annales de la Société entomologique de France (N.S.) 18(2): 181–198.

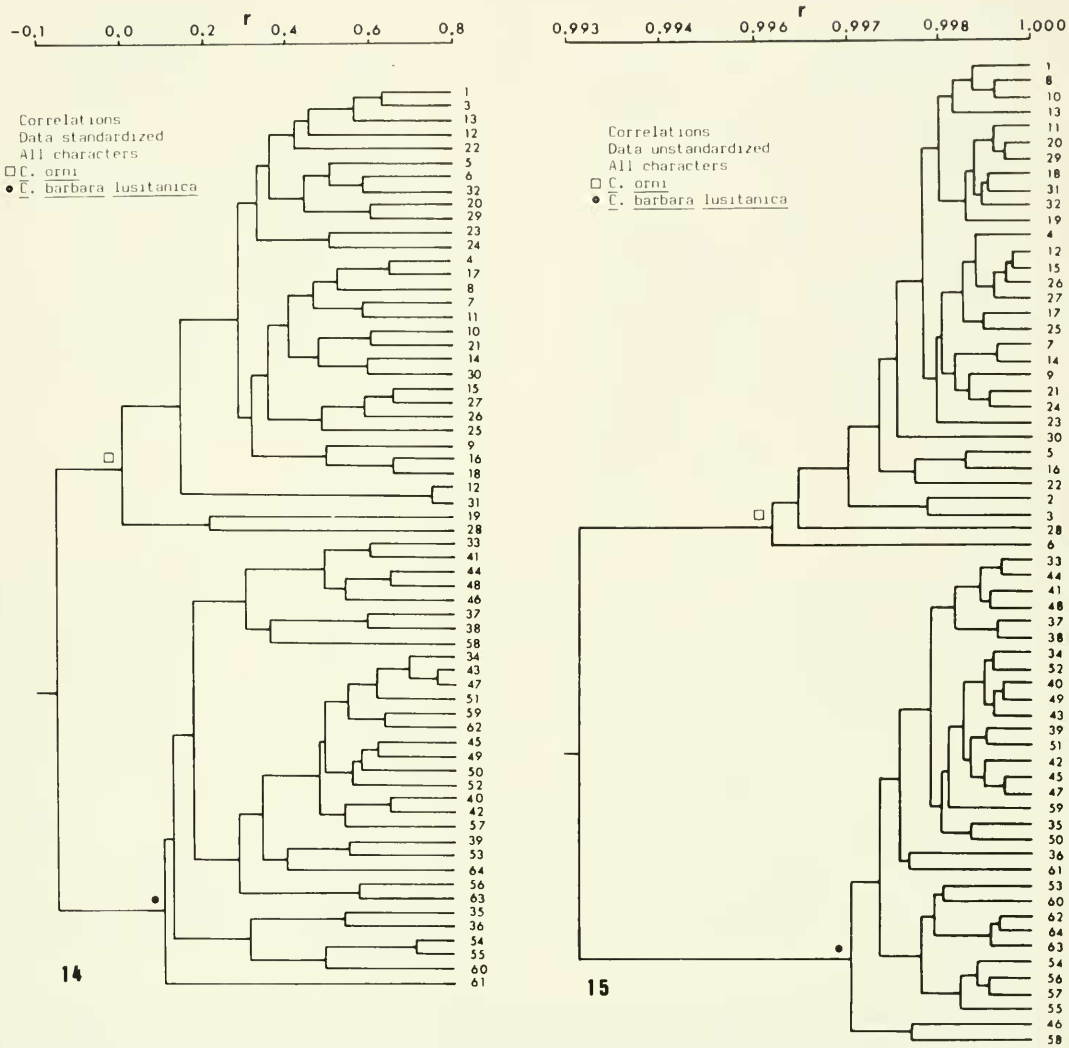
BOYCE, A. J. 1964. The value of some methods of numerical taxonomy with reference to hominoid classification. Pages 47–65 *in* V. H. Heywood and J. McNeill, eds., Phenetic and phylogenetic classification. Systematic Association Publication No. 6, London.

CHEETAM, A. H. 1968. Morphology and systematics of the bryozoan genus *Metrarabdotus*. Smithsonian Miscellaneous Collections 153(1): 1–121.

CLARIDGE, M. F., M. R. WILSON, AND J. S. SINGHRAO. 1979. The songs and calling sites of two European cicadas. Ecological Entomology 4: 225-229.

GIBSON, A. R., A. J. BAKER, AND A. MOEED. 1984. Morphometric variation in introduced populations of the common myna (*Acridotheres tristis*): an application of the jackknife to principal component analysis. Systematic Zoology 33(4): 408–421.

GÓMEZ-MENOR, O. J. 1957. Monografia de Cicadidos (Homoptera) de España. Academia de Madrid, Madrid.



Figs. 14–15. 14, correlation phenogram based on all 40 combined characters with standardized data; 15, correlation phenogram based on all 40 combined characters with unstandardized data.

HUMPHRIES, J. M., F. L. BOOKSTEIN, B. CHERNOFF, G. R. SMITH, R. L. ELDER, AND S. G. POSS. 1981. Multivariate discrimination by shape in relation to size. Systematic Zoology 30: 291–308.

JOLICOEUR, P., AND J. E. MOSIMANN. 1960. Size and shape variation in the painted turtle. A principal component analysis. Growth 24: 339–354.

MOSIMANN, J. E. 1970. Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. Journal of the American Statistical Association 65: 930–945.

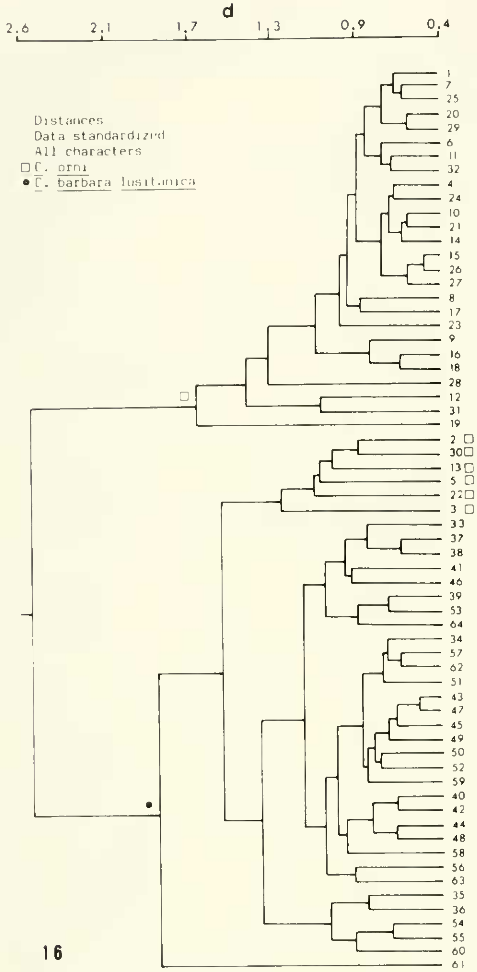
MYERS, J. G. 1928. The morphology of the Cicadidae (Homoptera). Proceedings of the Zoological Society of London 1928: 265–472.

OXNARD, C. E. 1978. One biologist's view of morphometrics. Annual Review of Ecology and Systematics 9: 219–241.

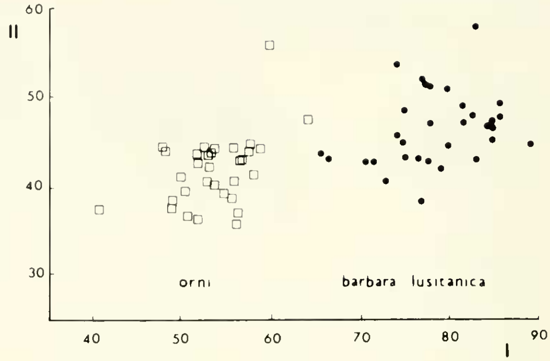
QUARTAU, J. A. 1983. A comparison of different numerical methods applied to the taxonomy of *Batracomorphus* Lewis (Insecta, Homoptera, Cicadellidae, Iassinae): ordination techniques. Journal of the Entomological Society of Southern Africa 46(1): 9–35.

______. 1987. Ordination vs. cluster analysis in the numerical taxonomy of some leafhoppers (Homoptera, Cicadellidae). Pages 207–215 *in* M. R. Wilson and L. R. Nault, eds., Proceedings of the 2nd International Workshop on Leafhoppers and Planthoppers of Economic Importance, C.I.E., London.

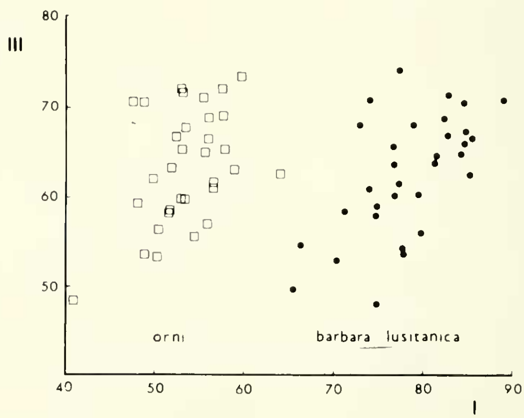
QUARTAU, J. A., AND R. G. DAVIES. 1983. A comparison of different numerical methods applied to the taxonomy of *Batracomorphus* Lewis (Insecta, Homoptera, Cicadellidae, Iassinae): cluster analysis. Revista de Biologia 12: 550–596.



16



17



18

Figs. 16–18. 16, distance phenogram based on all 40 combined characters with standardized data; 17, a two-dimensional view of the relationships among the 64 OTUs in a space determined by component I on x-axis (38.90%) and component II on y-axis (8.51%) of a principal component analysis of the matrix of correlations among all 40 characters with standardized data; 18, as in Fig. 17 but referred to component I on x-axis and III on y-axis (6.61%).

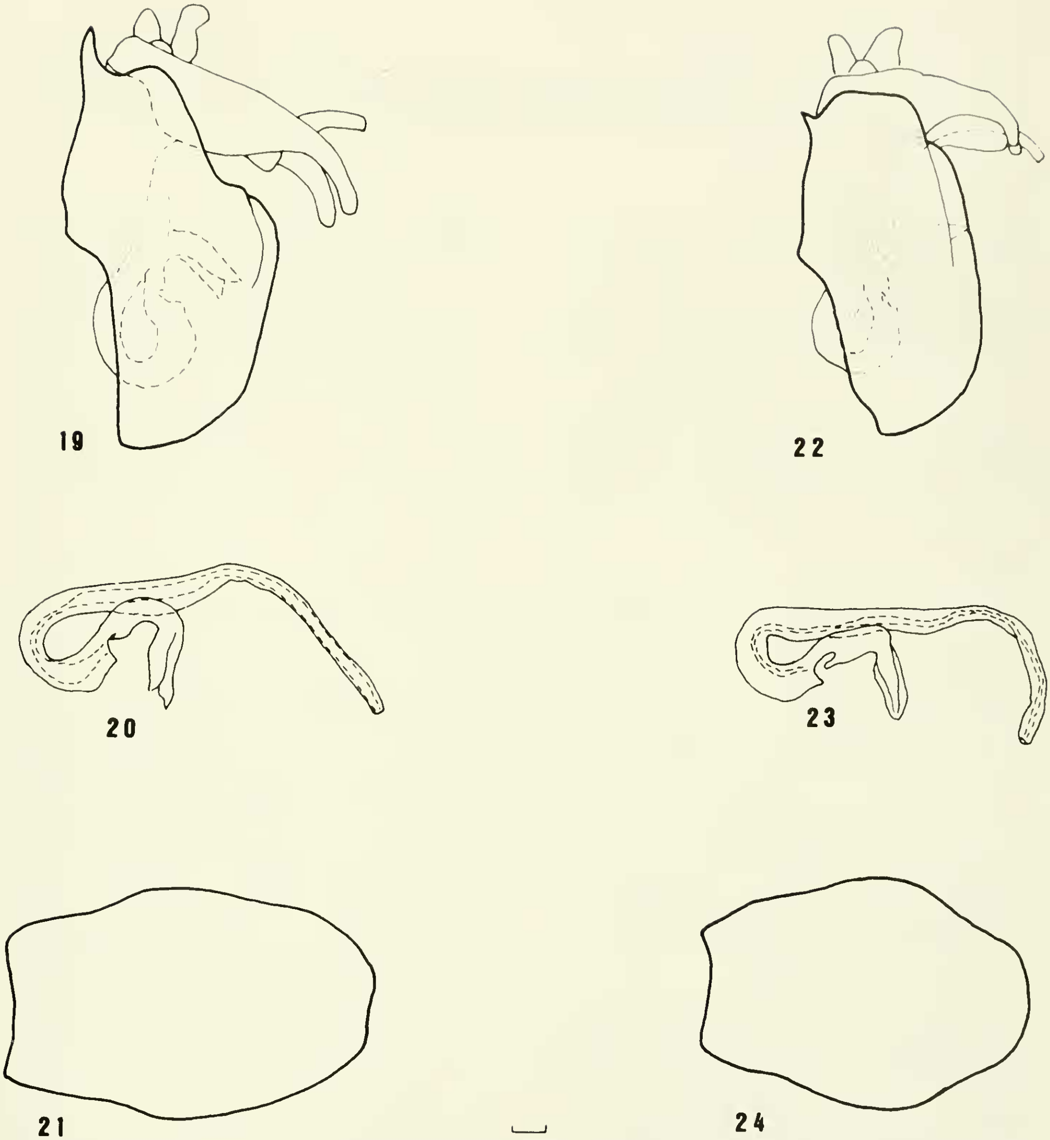
———. 1985. Character selection by information content in the numerical taxonomy of some male *Batracomorphus* (Homoptera: Cicadellidae). Zeitschrift für zoologische Systematik und Evolutionsforschung 23: 100–115.

QUARTAU, J. A., AND P. J. FONSECA. 1988. An annotated check-list of the species of cicadas known to occur in Portugal (Homoptera: Cicadoidea). Pages 367–375 *in* C. Vidano and A. Arzone, eds. Proceedings of the 6th Auchenorrhyncha Meeting, Turin.

SHEA, B. T. 1985. Bivariate and multivariate growth allometry: statistical and biological considerations. Journal of Zoology 206: 367–390.

SMITH, J. D. 1972. Systematics of the chiropteran family Marmoopidae. University of Kansas, Publications of the Museum of Natural History 56: 1–132.

TEMPLE, J. T. 1968. The Lower Llandovery (Silurian) brachiopods from Keisley, Westmorland. Palaeontographical Society 521: 1–58.



Figs. 19–24. Diagrams of the male genitalia of *Cicada orni* (Figs. 19–21) and of *C. barbara lusitanica* (Figs. 22–24): 19, 22, pygofer and tenth abdominal segment, lateral view; 20, 23, aedeagus, lateral view; 21, 24, eighth sternite or hypandrium, ventral view (scale = 0.5 mm).